Concept of Operation and Data Requirements

DESDynI-L

Science Data Systems in the Decadal Survey Era Workshop, June 25-26, 2009

Steve Kempler
DESDynI-L Data Ground Systems Manager
NASA/GSFC
Contributors: Gary Alcott, Scott Luthcke, Chris Lynnes, Michael Urban, Bruce Vollmer

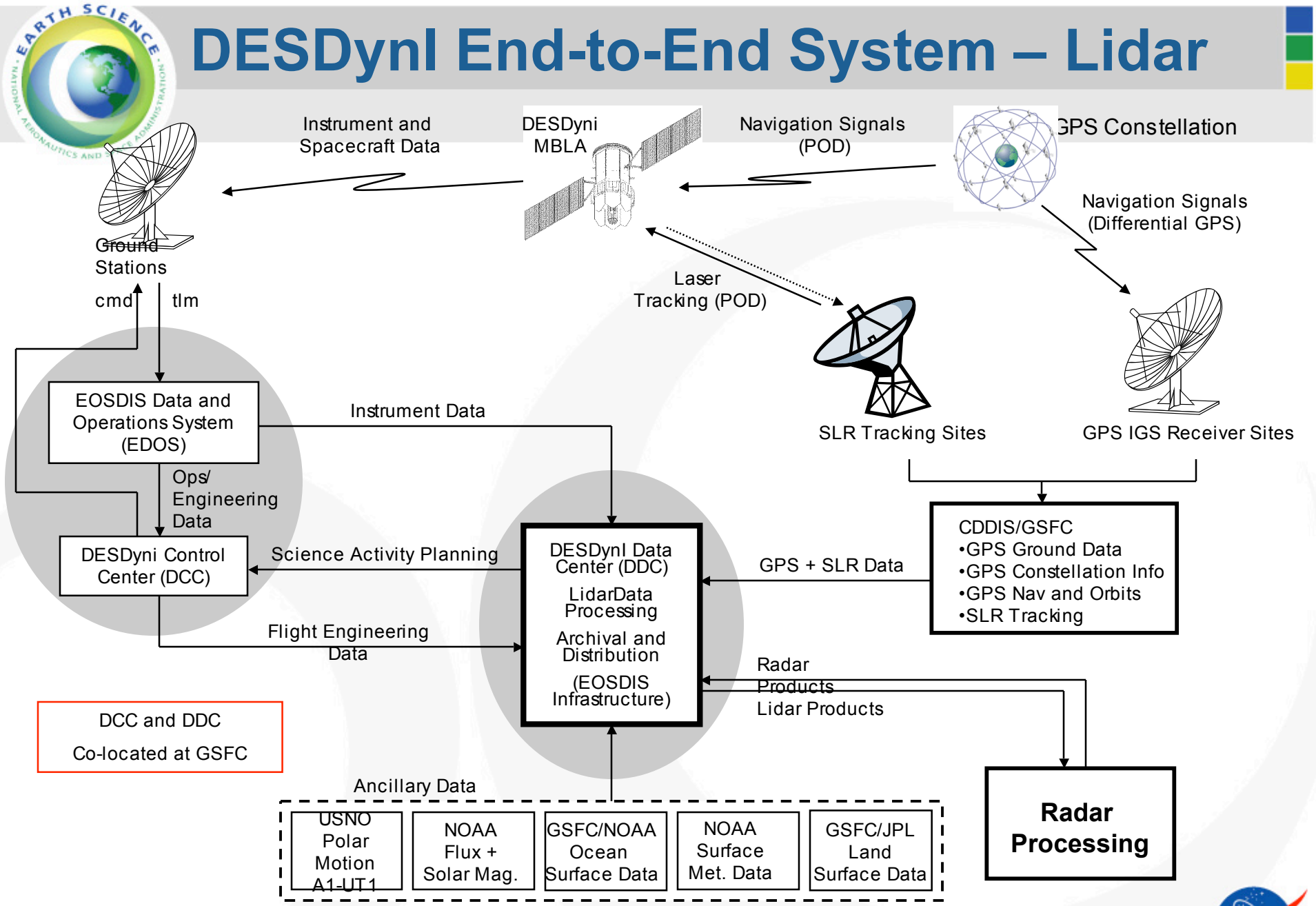
Overview of Presentation

- Mission Overview
- End-to-End Data Flow Diagram
- Key Driving Science Data System (SDS) Design Requirements
- Science Data Product Suite – Lidar
- SDS Architecture & Functional Diagram
- Algorithm-to-Production Software Process
- Data Storage and Archive Strategy
- Data Distribution Scenarios

DESDynI-L Mission Overview

- **DESDynI:** Deformation, Ecosystem Structure and Dynamics of Ice
- **Time-frame:** 2017 (2018) Launch, 5-year Mission, plus 3 years Data System Post Mission (mission ends, but data management does not)
- **Mission Objectives:**
  - Characterize the effects of changing climate on the carbon cycle
    - Determine biome and species level changes
  - Characterize the effects of changing land use on species habitats and carbon budget
- **Orbit:**
  - 400km, 91-day repeat for ecosystem sampling
- **Instrument:**
  - Multi-beam Lidar operating in the infrared

DESDynI End-to-End System – Lidar



Prepared by Scott B. Luthcke, NASA GSFC, Code 698, Jan. 22, 2009; Revised by M. Urban, SGT, June 5, 2009

Key Driving DDC Design Requirements

- Mission life: **5 Years**
- Number of input data sources: **9**
- Estimated total volume to archive: **345 TB**
- Data levels to process: **L1 - L3**
- Processing rates: **Start with 2X forward; 3X reprocess. Add 3X reprocessing each successive year of mission (X=processing a day of data in a day)**
- Reprocessing frequency: **4 months**
- Science Software Integration and Test (SSIT): **Algorithms written by science team, integrated and archived at data center**
- In addition:
  - Data ingest and metadata cataloging
  - Online data archive with online replication of critical data[1] and offsite backup
  - Data search and access
  - User support and outreach
  - Data processing as required
  - Project documentation
  - Lights-dim operations with 8x5 onsite support and off-hours automatic paging for remote support (8x7 for GPS data handling)

1 Critical data are data that are irrecoverable, difficult, or time-consuming to reproduce; typically raw data, ancillary, and level 3 gridded.

DESDynI Science Data Product Suite - Lidar

- **Level 0 and other instrument Products**
  - Uncompressed, depacketized and in engineering units
  - Engineering, GPS and Nav, StarTracker (3 trackers), Gyro data

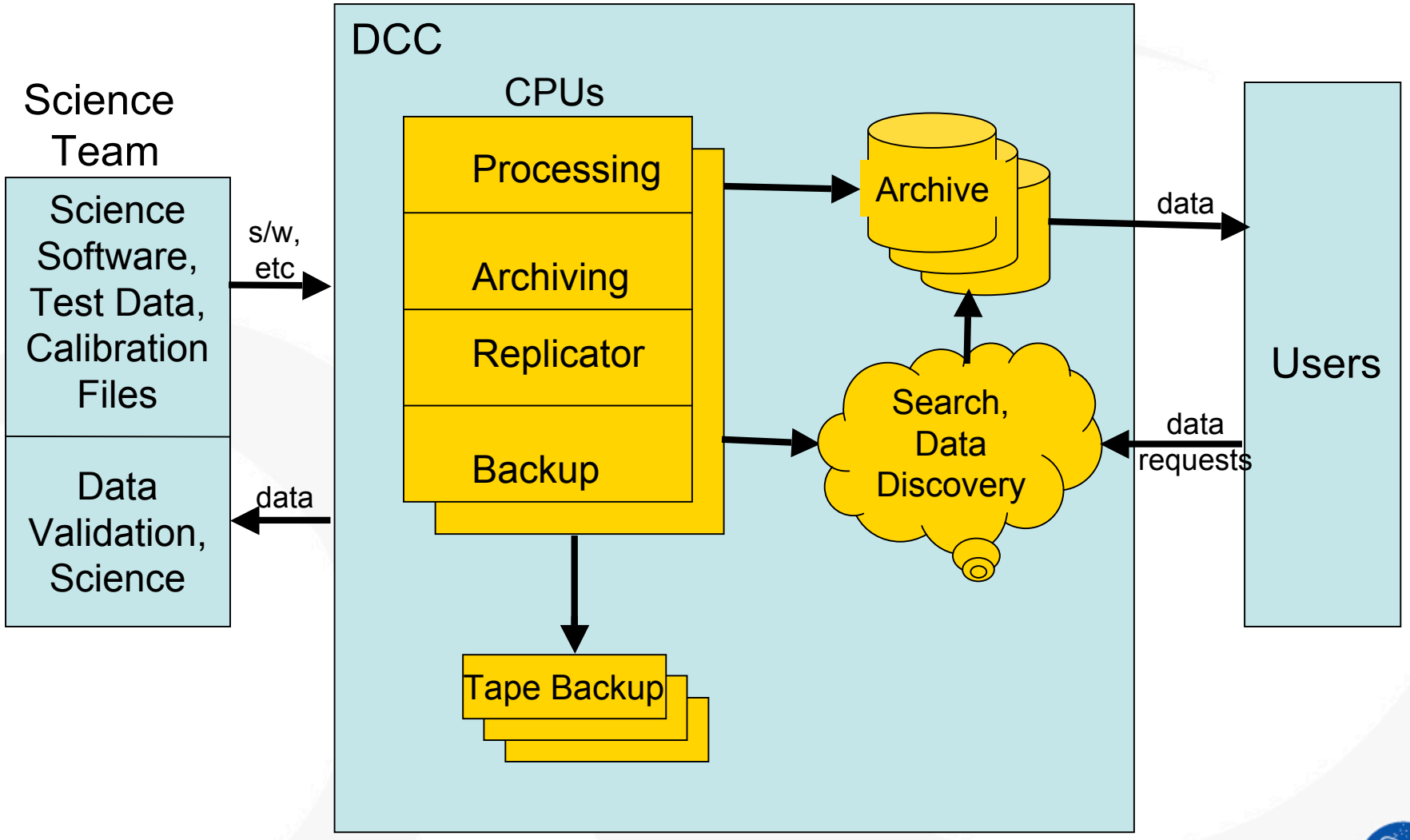
- **Level 1 Lidar Instrument Products**
  - Precision Orbit Determination (POD)
  - Precision Attitude Determination (PAD)
  - Integrated Residual Analysis (IRA) for pointing, ranging and timing calibration
  - Calibrated waveforms and ranges

- **Level 2/3 Science Products**
  - Geolocated Waveforms
  - Gridded Science Products (5): seasonally, yearly, mission

- **Ancillary file data acquired each day**
  - Polar Motion/A1-UT1 from USNO
  - Solar Mag. Flux from NOAA
  - Surface Met. Data from NOAA
  - Ocean topo. and wave from NOAA
  - SLR data from CDDIS
  - IGS GPS ground data
  - IGS GPS orbit data

- **Static ancillary data (for the entire mission):**
  - Planetary Ephemeris (2 files)
  - Gravity Fields (3 files)
  - Time variable gravity (3 ATGRAV)
  - BIH Tables (5 files)
  - Mean Sea Surface (3 files)
  - Tides supporting files
  - DEMs, 100 m resolution (3 files)
  - Misc. files

DESDynI-L Data Center (DCC) Architecture & Functional Diagram



Algorithm-to-Production Software Process

- 3 years prior to launch - Begin capabilities/documentation, interface discussions, algorithm requirements meetings
- 2 years prior to launch - Begin documenting interface agreements, collecting algorithm and data documentation
- 1.5 years prior to launch - Procure testbed hardware for initial algorithm testing; Refine science software integration and test procedures
- 1 year prior to launch - Procure full operations system hardware, system integration, and science software integration
- 8 months prior to launch - Begin system testing with 'final' science software
- 4 months prior to launch - Begin end-end testing with REALLY final science software; Continue to run test data through system
- October, 2017 (2018) - Launch 
- With first data - Support launch and early checkout
- During life of mission – Integrate improved algorithms, process/reprocess data
- During post mission - Integrate improved algorithms, reprocess data

Data Storage and Archive Strategy

- Reuse existing open-source operational data processing and archive systems, thus removing new system development costs; No COTS software = No compatibility issues and upgrade maintenance
- Examine co-locating DESDynI data science management system with other science data systems to leverage/share existing infrastructure and personnel
- Purchase hardware yearly, the year before it is needed, to take advantage of lowering hardware costs
- Maximize use of data standards (e.g., metadata, format, software, others)
- Implement and shake out basic required data system; Once data arrive and are better understood, seek to integrate existing tools that help prepare data for science research
- Data system to be implemented in close coordination with the science team
- All technologies needed to meet the requirements exist.
  - With that said, plan to utilize SAN archive
  - Host-attached RAID is an adequate fallback if SAN prices do not drop sufficiently

Data Distribution Scenarios

To accommodate the varied skills, knowledge, and needs of data and information users, a variety of data distribution scenarios will be available:

- Knowledgeable Users (e.g., Researchers) who know the specific data they are interested in
  - Perform Google-like web-based searches for data
  - Perform web-based searches by time or location
  - Data on-line, thus access via ftp (no data ordering necessary)
- Knowledgeable Users (e.g., Researchers) who are looking for specific known geophysical signatures
  - Explore and discover data with browse visualizations
  - Download desirable data when found
- Novice Users (e.g., could be applications researchers, students)
  - Search for known events at/or specific locations
  - Provide data translation information relevant to application
  - Download data/images using web mapping services standards
- Multi-sensor data users
  - Seemlessly provide access to local and remote datasets using OpenDAP or the latest best technology
  - Perform multi-site data access compatible to OGC standards